

Genome-wide Association Studies

Jing Hua Zhao, MRC Epidemiology Unit, Cambridge, UK

Owing to advances in genotyping and sequencing technologies and the successes of a number of international collaborative projects, there is a considerable interest in genetic study of complex traits which include common diseases and other quantitative measurements. The current focus has been genome-wide association studies (GWAS) which involve a large number of individuals each with data from GeneChip containing 10,000~1,000,000 single nucleotide polymorphisms (SNPs), the most abundant genetic variants in human genome.

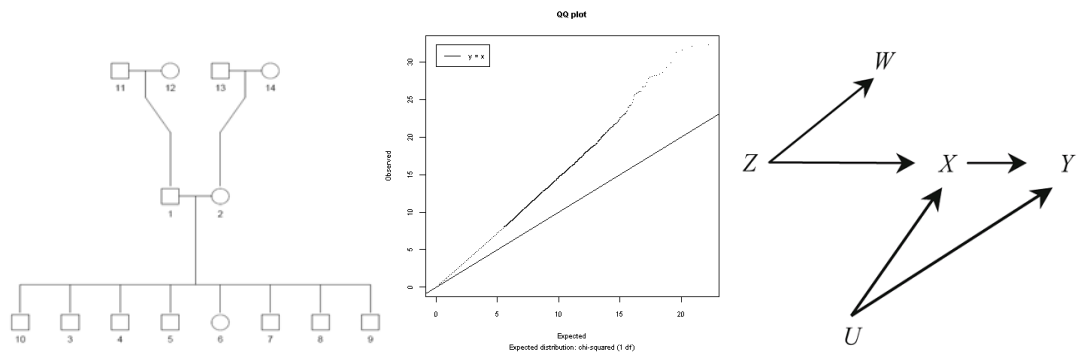
The tutorial begins with some fundamental concepts and methods which naturally lead to GWAS, to be followed by their study designs and statistical analysis. This ranges from quality controls and descriptive analysis to assessing genotype-phenotype relationship and inference of pathways. Genotype imputation, meta-analysis and graphical presentation will be illustrated with case studies involving population-based and family-based samples. Publicly available projects, such as HapMap and dbGaP, will also be introduced. The tutorial and materials will enable the participants become familiar with the computational and statistical problems, popular specialized software and relevant packages in R, e.g., *genetics*, *haplo.stats*, *gap*, *kinship*, as with packages designed for GWAS such as *GenABEL*, *snpMatrix* and *SNPassoc*.

The tutorial is appropriate for researchers with basic knowledge in statistics and computing and wish to get involved with genetic data analysis in humans while generating interest to researchers in plant and animal sciences. However, it will also be useful to professionals and researchers actively engaged in analysis of genetic data and/or development of computational tools in R or other environments

The tutorial is derived from considerable theoretical and practical work in statistical genetics and genetic epidemiology, especially from design and analysis for GWAS in several large epidemiological cohorts and consortium work.

References

- Altshuler D, MJ Daly, ES Lander. Genetic mapping in human disease. *Science* 322:881-888, 2008
- Donnelly P. Progress and challenges in genome-wide association studies in humans. *Nature* 456:728-731, 2008
- Elston RC, M Ann Spence. Advances in statistical human genetics over the last 25 years. *Statistics in Medicine* 25: 3049-3080, 2006
- Pearson TA, TA Manolio. How to interpret a genome-wide association study. *JAMA* 299:1335-1344, 2008
- Zhao JH, Q Tan. Integrated analysis of genetic data with R. *Human Genomics*. 2:258-265, 2006



Schedule (Tuesday 7/7/2009, 9:00 am – 12:30 pm)

9:00 am	—	9:45 am	Overview
9:45 am	—	10:30 am	Association Analysis
10:30 am	—	11:00 am	Break
11:00 am	—	11:45 am	Meta-Analysis
11:45 am	—	12:30 pm	Other topics

The tutorial will be lecture only and laptop is not required. It has a web site:

www.mrc-epid.cam.ac.uk/~jinghua.zhao/notes/useR2009